

【公報種別】特許法第 17 条の 2 の規定による補正の掲載  
 【部門区分】第 6 部門第 3 区分  
 【発行日】令和 3 年 9 月 2 日 (2021.9.2)

【公表番号】特表 2020-532010 (P2020-532010A)  
 【公表日】令和 2 年 11 月 5 日 (2020.11.5)  
 【年通号数】公開・登録公報 2020-045  
 【出願番号】特願 2020-511746 (P2020-511746)  
 【国際特許分類】

G 0 6 N 20/00 (2019.01)

【 F I 】

G 0 6 N 20/00

【手続補正書】

【提出日】令和 3 年 7 月 21 日 (2021.7.21)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

異常値を検知するシステムであって、

複数の物理的エンティティ記録を含むデータベースに対するアクセスを有する通信モジュールであって、各物理的エンティティ記録は、少なくとも 1 つの数値属性 (13、14) の物理データ値と、1 つ以上の名義属性の値に関するパーティション指定値 (11、12) とを含む、通信モジュールと、

プロセッサ、有形不揮発性メモリ、前記プロセッサに命令する、前記メモリ上に存在するプログラムコードを備えるコンピューティングデバイスと、  
 を備え、

前記通信モジュールは、前記データベースに対するアクセスを前記コンピューティングデバイスに提供するように構成され、

前記コンピューティングデバイスは、前記複数の物理的エンティティ記録の各々について異常スコアを計算する方法を実行するように構成され、前記方法は以下のステップ、

(a) 前記通信モジュールを介して前記複数の物理的エンティティ記録を索出し、任意選択で前記複数の物理的エンティティ記録をパーティション化のために準備するステップと、

(b) 前記複数の物理的エンティティ記録内に存在するパーティション指定値 (11、12) の各個別の組み合わせにパーティションを関連付けることによって、前記複数の物理的エンティティ記録をパーティション化し、前記パーティションに従って前記複数の物理的エンティティ記録をグループ分けするステップと、

(c) ステップ (b) で取得された前記パーティションごとに、該パーティションに属する前記物理的エンティティ記録の前記物理データ値に対して教師なし異常検知アルゴリズムをトレーニングして、前記パーティションごとにトレーニングされた異常検知モデルを取得するステップと、

(d) 前記複数の物理的エンティティ記録に属する物理的エンティティ記録ごとに、前記物理的エンティティ記録が属する前記パーティションに関連付けられた前記トレーニングされた異常検知モデルを用いて前記異常スコア (17) を計算するステップと、

(e) ステップ (d) において計算されたその関連する異常スコア (17) を追加することによって、前記通信モジュールを介して、前記データベース内の各物理的エンティティ記

録を更新し、及び／又は、前記通信モジュールを介して、前記データベース内の前記パーティションごとの前記トレーニングされた異常検知モデルの各々を記憶するステップと、のうち、(a)～(d)のステップ又は(a)～(e)のステップを含む、システム。

【請求項2】

前記教師なし異常検知アルゴリズムは、iForest若しくはSCiForest若しくはiNNE等の隔離ベース異常検知アルゴリズムか、又は、ORCA若しくは局所外れ値因子法等の非隔離ベース異常検知アルゴリズムのいずれかに関し、前記パーティション指定値(11、12)は、2つの名義属性の値に関する、請求項1に記載のシステム。

【請求項3】

前記物理データ値は画像に関し、前記パーティション指定値(11、12)は、EXIF情報、ファイル日付、又はファイル名等の前記画像に対するメタデータとして記憶された名義属性の値に関する、請求項1又は2に記載のシステム。

【請求項4】

前記複数の物理的エンティティ記録は、ステップ(a)におけるパーティション化に向けて準備され、該準備することは、前記複数の物理的エンティティ記録を更新することを含み、該更新することは、いずれの名義属性をパーティション指定値(11、12)としてステップ(b)以降に用いるかを、自動的に、指定することを含む、請求項1～3のいずれか一項に記載のシステム。

【請求項5】

前記通信モジュールはユーザーから入力を受信するように構成され、ステップ(a)における前記準備することの一部としての前記更新することは、前記通信モジュールを介して前記ユーザーから前記入力を受信することを含み、前記ユーザーからの前記入力は、いずれの名義属性をパーティション指定値(11、12)としてステップ(b)以降に用いるかの前記指定を含む、請求項4に記載のシステム。

【請求項6】

前記複数の物理的エンティティ記録の前記更新は、少なくとも1つの数値属性を新規作成される名義属性に、自動的に、変換することを含み、該変換することは、少なくとも1つの名義ラベルを少なくとも1つの数値区間に関連付けることに関連し、いずれの名義属性をパーティション指定値(11、12)としてステップ(b)以降に用いるかの前記指定は前記新規作成された名義属性を含む、請求項4又は5に記載のシステム。

【請求項7】

前記データベースは、前記複数の物理的エンティティ記録に含まれる複数のトレーニング関連の物理的エンティティ記録に関連するトレーニングデータを含み、該トレーニングデータは、前記トレーニング関連の物理的エンティティ記録が、詐欺に関連している等の規定の異常関連基準を遵守しているか否かを示す複数のラベルを含み、前記コンピューティングデバイスは、ステップ(e)の後に以下の追加ステップ(f)～(j)、

(f)ステップ(d)において計算されたその関連する異常スコア(17)を追加することによって、各物理的エンティティ記録を更新するステップと、

(g)前記通信モジュールを介して前記トレーニングデータを索出するステップと、

(h)ステップ(g)において受信された前記トレーニングデータに基づいて、ステップ(f)において取得された前記更新された物理的エンティティ記録に対して教師あり分類アルゴリズムをトレーニングし、トレーニングされた教師あり分類モデルを取得するステップと、

(i)前記複数の物理的エンティティ記録に属する物理的エンティティ記録ごとに、予測スコアを計算するステップであって、該予測スコアは、前記トレーニングされた教師あり分類モデルを用いて、前記物理的エンティティ記録が前記規定の異常関連基準をどの程度遵守しているかを示す、ステップと、

(j)ステップ(i)において計算されたその関連する予測スコアを追加することによって、前記通信モジュールを介して、前記データベース内の各物理的エンティティ記録を更新し、及び／又は、前記通信モジュールを介して、前記データベース内の前記トレーニン

グされた教師あり分類モデルを記憶するステップと、  
のうち、(f) ~ (i) のステップ又は (f) ~ (j) のステップを実行するように構成される、請求項1~6のいずれか一項に記載のシステム。

【請求項 8】

物理的エンティティ記録ごとに、ステップ(f)における更新することは、前記パーティション指定値(11、12)を、前記物理的エンティティ記録と関連付けられた前記異常スコア(17)に交換することに関する、請求項7に記載のシステム。

【請求項 9】

ステップ(h)においてトレーニングされ、ステップ(i)において適用された前記教師あり分類アルゴリズムは、ロジスティック回帰、又はCART決定木、又はランダムフォレスト、又は線形カーネルを用いたSVM、又は放射基底関数を用いたSVMに関する、請求項7又は8に記載のシステム。

【請求項 10】

ステップ(h)における前記教師あり分類アルゴリズムの前記トレーニングは、ホールドアウト、k分割交差検証のいずれかのストラテジーに従って、前記トレーニング関連の物理的エンティティ記録及び関連したトレーニングデータを分割することを含む、請求項7~9のいずれか一項に記載のシステム。

【請求項 11】

前記通信モジュールは、ユーザーから入力を受信するように更に構成され、該システムは、以下の更なるステップ、

(01) 前記ユーザーから前記通信モジュールを介してユーザー提供の物理的エンティティ記録を受信するステップであって、該ユーザー提供の物理的エンティティ記録は、少なくとも1つの数値属性(13、14)の物理データ値と、1つ以上の名義属性の値に関するパーティション指定値(11、12)とを含む、ステップと、

(02) スコアリングのために前記ユーザー提供の物理的エンティティ記録を準備するステップであって、該準備することは、パーティション指定値(11、12)に対応する1つ以上の名義属性を選択することを含む、ステップと、

(03) ステップ(02)で選択された前記名義属性に対応する前記トレーニングされた異常検知モデルを、前記通信モジュールを介して前記データベースから索出するステップと、

(04) ステップ(03)で索出された、前記トレーニングされた異常検知モデルを用いて、前記ユーザー提供の物理的エンティティ記録の前記異常スコア(17)を計算するステップと、

を実行するように構成される、請求項1~10のいずれか一項に記載のシステム。

【請求項 12】

前記データベースは、少なくとも1つの教師あり分類モデルを更に含み、該システムは、ステップ(04)の後に以下の追加ステップ(05)~(07)、

(05) 前記ユーザー提供の物理的エンティティ記録を、ステップ(04)で計算されたその関連する異常スコア(17)を追加することによって更新するステップと、

(06) ステップ(02)で選択された前記名義属性に対応する前記トレーニングされた教師あり分類モデルを、前記通信モジュールを介して前記データベースから索出するステップと、

(07) ステップ(06)で索出された、前記トレーニングされた教師あり分類モデルを用いて、前記ユーザー提供の物理的エンティティ記録の前記予測スコアを計算するステップと、

を実行するように構成される、請求項11に記載のシステム。

【請求項 13】

前記通信モジュールは、オペレーターに対してアラート、又は視覚的若しくは音響アラートを生成するように更に構成され、該システムは、ステップ(04)で計算された前記異常スコア(17)、又はステップ(07)で計算された前記予測スコアを規定のアラート値と

比較するように更に構成され、前記オペレーターに対する前記アラートは、前記比較することによって、前記規定のアラート値が一致したか、又は該アラート値を超えた場合に生成される、請求項11又は12に記載のシステム。

【請求項 14】

複数の物理的エンティティ記録に関して異常値を検知するコンピューター実施方法であって、各物理的エンティティ記録は、少なくとも1つの数値属性（13、14）の物理データ値と、1つ以上の名義属性の値に関するパーティション指定値（11、12）とを含み、該方法は、

（a）データベースから前記複数の物理的エンティティ記録を取得し、任意選択で前記複数の物理的エンティティ記録をパーティション化のために準備するステップと、

（b）前記複数の物理的エンティティ記録内に存在するパーティション指定値（11、12）の各個別の組み合わせにパーティションを関連付けることによって、前記複数の物理的エンティティ記録をパーティション化し、前記パーティションに従って前記複数の物理的エンティティ記録をグループ分けするステップと、

（c）前記パーティション化ステップで取得された前記パーティションごとに、該パーティションに属する前記物理的エンティティ記録の前記物理データ値に対して教師なし異常検知アルゴリズムをトレーニングし、前記パーティションごとにトレーニングされた異常検知モデルを取得するステップと、

（d）前記複数の物理的エンティティ記録に属する物理的エンティティ記録ごとに、前記物理的エンティティ記録が属する前記パーティションに関連付けられた前記トレーニングされた異常検知モデルを用いて異常スコア（17）を計算するステップと、

（e）前記計算ステップにおいて計算されたその関連する異常スコア（17）を追加することによって、各物理的エンティティ記録を更新し、及び／又は、前記データベース内の前記パーティションごとの前記トレーニングされた異常検知モデルの各々を記憶するステップと、

のうち、（a）～（d）のステップ又は（a）～（e）のステップを含む、方法。